

MCDONALD, S., QUINN, F., VIEIRA, R., O'BRIEN, N., WHITE, M., JOHNSTON, D.W. and SNIEHOTTA, F.F. 2017. The state of the art and future opportunities for using longitudinal n-of-1 methods in health behaviour research: a systematic literature overview. *Health psychology review* [online], 11(4), pages 307-323. Available from: <https://doi.org/10.1080/17437199.2017.1316672>

The state of the art and future opportunities for using longitudinal n-of-1 methods in health behaviour research: a systematic literature overview.

MCDONALD, S., QUINN, F., VIEIRA, R., O'BRIEN, N., WHITE, M., JOHNSTON, D.W., SNIEHOTTA, F.F.

2017



The state of the art and future opportunities for using longitudinal n-of-1 methods in health behaviour research: a systematic literature overview

Suzanne McDonald^a, Francis Quinn^b, Rute Vieira^a, Nicola O'Brien^a, Martin White^c, Derek W. Johnston^d and Falko F. Sniehotta^e

^aInstitute of Health & Society, Newcastle University, Newcastle upon Tyne, UK; ^bSchool of Applied Social Studies, The Robert Gordon University, Aberdeen, UK; ^cUKCRC Centre for Diet and Activity Research (CEDAR), MRC Epidemiology Unit, School of Clinical Medicine, University of Cambridge, Cambridge, UK; ^dSchool of Psychology, University of Aberdeen, Aberdeen, UK; ^eFuse, UKCRC Centre for Translational Research in Public Health, Newcastle University, Newcastle upon Tyne, UK

ABSTRACT

n-of-1 studies test hypotheses within individuals based on repeated measurement of variables within the individual over time. Intra-individual effects may differ from those found in between-participant studies. Using examples from a systematic review of n-of-1 studies in health behaviour research, this article provides a state of the art overview of the use of n-of-1 methods, organised according to key methodological considerations related to n-of-1 design and analysis, and describes future challenges and opportunities. A comprehensive search strategy (PROSPERO:CRD42014007258) was used to identify articles published between 2000 and 2016, reporting observational or interventional n-of-1 studies with health behaviour outcomes. Thirty-nine articles were identified which reported on n-of-1 observational designs and a range of n-of-1 interventional designs, including AB, ABA, ABABA, alternating treatments, n-of-1 randomised controlled trial, multiple baseline and changing criterion designs. Behaviours measured included treatment adherence, physical activity, drug/alcohol use, sleep, smoking and eating behaviour. Descriptive, visual or statistical analyses were used. We identify scope and opportunities for using n-of-1 methods to answer key questions in health behaviour research. n-of-1 methods provide the tools needed to help advance theoretical knowledge and personalise/tailor health behaviour interventions to individuals.

ARTICLE HISTORY


Received 26 September 2016
Accepted 2 April 2017

KEYWORDS

n-of-1 methods; health behaviour; single-case; idiographic methods

Health behaviours differ between individuals. Some people are more physically active than others, some binge drink, smoke or eat five portions of fruit and vegetables per day, whereas others do not. These inter-individual differences are often the object of behavioural and psychological research, for example, research establishing covariates of inter-individual differences or evaluating the effects of interventions by randomly allocating groups of individuals to different treatments and comparing outcomes. In health psychology and behavioural medicine, between-participant designs are frequently used to describe behavioural phenomena, identify relationships between variables and evaluate effects of behavioural interventions. However, conclusions from

CONTACT Suzanne McDonald  suzanne.mcdonald@ncl.ac.uk

 Supplemental data for this article can be accessed at doi:[10.1080/17437199.2017.1316672](https://doi.org/10.1080/17437199.2017.1316672).

© 2017 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

between-participant studies may not apply to individuals. For example, trials may find that an intervention is effective for the group receiving it on average but may bring no benefit or even harmful effects to some individuals. Between-participant observational cohort designs are often used to draw conclusions about behaviour over time (e.g., physical activity levels decline as individuals get older; Scholes & Mindell, 2013) but individual trajectories of behaviour over time may not follow the average pattern and may instead vary considerably between individuals (Johnston & Johnston, 2013). Research methods are needed that can enrich the description of phenomena, identification of relationships between variables and evaluation of effects of interventions within individuals.

n-of-1 (also known as 'single-case') methods involve the repeated measurement of an individual (or individual unit, for example, one family or one hospital) over time, allowing conclusions to be drawn about the individual. n-of-1 methods are recognised and recommended by the UK Medical Research Council as useful for testing theory and interventions (Medical Research Council, 2008). They can reveal how health behaviours change over time within individuals, which is not well represented in findings from group-based studies. Investigators can use n-of-1 methods to describe intra-individual patterns of behaviour over time, examine relationships between potential predictors of behaviour over time and to identify individual response to interventions. Using methods which capture intra-individual variability in health behaviour is important for selecting interventions that are effective for changing an individual's health behaviour. n-of-1 methods can also be used to personalise interventions to individuals (McDonald, Araujo-Soares, & Sniehotta, 2016a). For example, identifying the unique predictors of an individual's behaviour can lead to the design of a personalised intervention which targets those predictors (O'Brien, Philpott-Morgan, & Dixon, 2016).

Where n-of-1 protocols are used with more than one participant, it is possible to test how generalisable conclusions made for individual cases are, or even to identify between-participant moderators of within-participant processes. This can be achieved through random effects meta-analysis, the summary measures approach or multilevel modelling (Araujo, Julious, & Senn, 2016). For example, if a series of n-of-1 studies share a sufficiently common methodology, multilevel methods can be used to aggregate n-of-1 data to determine the overall statistical significance of a series of studies (e.g., Sniehotta, Pesseau, Hobbs, & Araujo-Soares, 2012). Multilevel methods have been in use for some time and are well-documented (Hox, Moerbeek, & van de Schoot, 2010). The unique feature of n-of-1 studies is the ability to answer research questions and come to conclusions for an individual case.

n-of-1 methods have a long history of use in special education (Moeller, Dattilo, & Rusch, 2015), medicine (Gabler, Duan, Vohra, & Kravitz, 2011; Punja et al., 2016), neuropsychological rehabilitation (Perdices & Tate, 2009), psychotherapy (Norcross & Wampold, 2011) and various subdivisions of psychology (Smith, 2012a). The definition of n-of-1 methods and the characteristics of baseline measurement, sampling, outcomes and methods of analysis differ widely across the various fields in which n-of-1 methods have been used. For example, pharmacological studies tend to adopt n-of-1 designs which rely on the withdrawal of interventions. Intervention withdrawal is not always possible with interventions used to change health behaviours in health psychology and behavioural medicine. With the development of affordable and ubiquitous means of measuring individual behaviours outside of the laboratory (e.g., real-time self-report diaries, accelerometers and geo-sensors in mobile phones), there has been a considerable increase in the use of n-of-1 methods as a viable method to study health behaviour (McDonald & Davidson, 2016). However, the extent of and purpose for which n-of-1 methods are used in health behaviour research is unknown.

The aim of this article is to provide an introduction to the use of n-of-1 methods in health behaviour research and a description of the range of n-of-1 designs and methods of analysis that can be used to study health behaviour. It describes the state of the art, using examples from a systematic review of n-of-1 studies which have been applied to study or change health behaviour, and highlights considerable opportunities and challenges for future single-case health behaviour research.

Method

Systematic review search strategy and study eligibility

The databases PsycINFO, Embase and MEDLINE were searched using a comprehensive search strategy (see Supplemental File 1) on 13th May 2016. A request for published articles meeting inclusion criteria was sent to members of the European Health Psychology Society (EHPS), EHPS n-of-1 special interest group, UK Society for Behavioural Medicine, International Society of Behavioural Medicine and Society of Behavioural Medicine. The systematic review adhered to a registered protocol (McDonald, Quinn, Hobbs, White, & Sniehotta, 2014) and followed systematic review guidelines (see PRISMA checklist in Supplemental File 4). Observational or interventional n-of-1 studies reporting data analysis and conclusions at the individual level were included. Advances in technology and sampling procedures that enabled reliable and valid measurement of behaviours were starting to appear at the turn of the twenty-first century (Dallery, Kurti, & Erb, 2015) and are important for the conduct of rigorous n-of-1 research. For example, Ecological Momentary Assessment (EMA), which, using technology, involves real-time sampling of variables in a participant's natural environment, was introduced during this period and has revolutionised the field of psychological and behavioural measurement (Stone, 2000). To provide a contemporary picture of the use of n-of-1 methods in health behaviour research, studies published from 1st January 2000 were included in the systematic review. Studies with interventional n-of-1 designs were included if the intervention targeted a health behaviour. Studies with participants of any age, gender or health status were included. Studies that measured health behaviour (e.g., physical activity, food intake, smoking, alcohol consumption, drug use, adherence to treatment, UV exposure) as outcomes, using self-report or objective methods at two or more time points, were included. A detailed protocol describing the search strategy, study eligibility criteria, screening procedure, reliability assessment, data extraction and analysis is provided in Supplemental File 1.

Data synthesis

Included articles were synthesised descriptively according to their design and method of analysis. Articles were used as illustrative examples of the key considerations involved in the design and analysis of n-of-1 research. For example, strengths and challenges, and, where possible, good practice, related to each design were highlighted using examples.

Results

The systematic review identified 39 articles using n-of-1 methods to study or change health behaviours (see Supplemental Figure 1 for the number of publications included and excluded at each stage). A variety of n-of-1 designs were adopted including observational ($n = 2$) and interventional designs such as AB ($n = 9$), ABA ($n = 12$) ABABA ($n = 1$), alternating treatments ($n = 3$), n-of-1 randomised controlled trials ($n = 3$), multiple baseline ($n = 8$) and changing criterion ($n = 1$) designs. n-of-1 studies included patients and healthy volunteers ranging from 2 to 89 years of age and assessed one or more health behaviours including treatment adherence ($n = 16$), physical activity ($n = 14$), drug ($n = 6$) and alcohol ($n = 3$) use, sleep ($n = 2$), smoking ($n = 4$) and eating ($n = 1$) behaviour. Most studies used visual analysis ($n = 21$) to evaluate n-of-1 data, whilst others used descriptive ($n = 7$) or statistical analysis ($n = 11$) (see Supplemental Table 1 for further characteristics of included n-of-1 studies).

Designing n-of-1 research

The selection of a specific n-of-1 design is strongly influenced by the research question of interest. Three main categories of research questions can be explored using n-of-1 methods: (1) how do behaviours change over time within individuals? (i.e., the description of behaviours); (2) what is the

relationship between behaviours and various correlates over time within individuals? (i.e., the description of relationships); and (3) how do interventions impact on behaviours over time within individuals? (i.e., the description of the impact of interventions). Studies that aim to describe behaviours or the relationship between behaviours and other correlates over time require an observational n-of-1 design, whereas studies that aim to evaluate interventions require an interventional n-of-1 design.

Observational n-of-1 designs

An observational n-of-1 design involves repeated measurements of behavioural outcomes over time within an individual. There is no manipulation of variables by the investigator (i.e., there is no controlled intervention implemented). Two studies in the review used an observational n-of-1 design (Hobbs, Dixon, Johnston, & Howie, 2013; Quinn, Johnston, & Johnston, 2013). The studies aimed to describe relationships between behaviour and predictors over time and test the validity of behavioural theories at the individual level. Both studies required participants to complete twice daily self-reported measurements of cognitions and behaviour either online or via a personal digital assistant for a period of between 6 and 11 weeks. Objective methods of physical activity measurement were also used. Hobbs et al. (2013) tested whether the Theory of Planned Behaviour (TPB; Ajzen, 1991) could predict three physical activity behaviours in six healthy volunteers. The TPB had variable intra and inter-individual predictive ability, predicting some but not all physical activity behaviours in five individuals and not predicting any physical activity behaviours in one individual. Quinn et al. (2013) tested an integrated model of disability (Johnston & Dixon, 2014), combining cognitive predictors from the TPB and pain to predict activity limitation in six women with chronic pain. Pain did not predict activity limitations in any of the participants, intention predicted in one participant, and perceived behavioural control predicted in the expected direction for one participant but in the opposite direction for another, contrary to previous findings. These observational n-of-1 studies demonstrate that the theories may not always be valid at the individual level.

Strengths of the observational n-of-1 design are that it can be useful for testing behavioural theory, studying naturally occurring phenomena and examining temporal patterns of behaviour or relationships between variables over time. Time lags within relationships can be identified (e.g., 'A' might lead to 'B', but not immediately). Information about individual differences in temporal relationships can inform the timing, duration and intensity of behaviour change interventions for individuals. However, the lack of experimental control and manipulation limits conclusions about causal relationships (Tate et al., 2016).

Interventional n-of-1 design

Interventional n-of-1 designs involve experimental manipulation to assess the effect of an intervention on behaviour. There are several types.

AB design. An AB design involves measurement of 'baseline' behaviour (phase A) before an intervention is implemented, and after the intervention is introduced (phase B). Nine studies in the review used an AB design to evaluate the impact of interventions on treatment adherence (Gonzalez et al., 2010; Gorski, Slifer, Townsend, Kelly-Suttka, & Amari, 2005; Piven & Duran, 2014; Sather, Forbes, Starck, & Rovers, 2007; Sevic et al., 2005; Soroudi et al., 2008), smoking cessation (MacPherson, Collado, Ninemann, & Hoffman, 2016), physical activity (O'Brien et al., 2016) and drug use (Lee et al., 2014). For example, Gonzalez et al. (2010) evaluated the effect of 10–12 sessions of cognitive behavioural therapy for adherence in five diabetic patients by examining changes in adherence assessed by a Medication Event Monitoring System, a device which objectively records when pill bottles are opened, and a glucose monitor. The authors concluded that the intervention was effective in improving medication adherence in four participants and glucose monitoring in three participants.

An AB study design provides an opportunity to use information about behaviour collected during the baseline period to specify the design of an intervention for an individual case. For example, in one AB study identified in the review, predictors of physical activity behaviour in individuals with

osteoarthritis were identified for each individual during a 6-week baseline A phase and this information was used to develop a data-driven intervention delivered in a 6-week B phase (O'Brien et al., 2016). AB designs are often regarded as 'pre-experimental' because they do not control for external factors which may occur at the same time as the intervention (Kazdin, 2011). Therefore, in many cases it is not possible to rule out alternative explanations for changes in behaviour such as history (i.e., an event occurring at the same time as the intervention that has influence on the behaviour), maturation (i.e., processes within the individuals changing over time, for example, learning, getting tired, development over time) and regression to the mean. Indeed, some AB studies identified in the review acknowledged the possibility that changes observed in the behaviour may be related to other factors. For example, Soroudi et al. (2008) acknowledged the possibility that observed improvements in HIV medication adherence after cognitive behavioural therapy was delivered could be related to the simultaneous substance abuse counselling participants were receiving at their methadone clinic (history effect). To draw valid conclusions about the effect of an intervention, further temporal replications of the AB sequence are desirable.

ABA intervention withdrawal design. Withdrawal designs such as ABA designs provide more reliable evidence of the causal effects of an intervention because they reduce potential sources of bias. They involve monitoring the effects of implementing and then removing an intervention to determine whether behaviour changes in the expected direction. When an ABA intervention withdrawal design is used, behaviour is monitored at baseline, after the implementation of the intervention and after the completion of the intervention. Twelve studies identified in the review used an ABA design to evaluate the impact of interventions on treatment adherence (Cortina, Somers, Rohan, & Drotar, 2013; Daughters, Magidson, Schuster, & Safren, 2010; Gray, Janicke, Fennell, Driscoll, & Lawrence, 2011; Payne, Eaton, Mee, & Blount, 2013; Penica & Williams, 2008), physical activity (Casey, Mackay-Lyons, Connolly, Jennings, & Rasmussen, 2014; Lowe, Watanabe, Baracos, & Courneya, 2013), drug and alcohol use (Norberg, Perry, Mackenzie, & Copeland, 2014; Wright & Thompson, 2002), eating behaviour (Hill, Masuda, Moore, & Twohig, 2015), smoking (Banducci, Long, & MacPherson, 2015) and sleep behaviour (McCrae, Tierney, & McNamara, 2005). For example, McCrae et al. (2005) delivered a behavioural intervention designed to improve sleep quantity and quality in four caregivers. Self-reported sleep diaries were completed every morning during a 2-week A phase and a 4- to 8-week B phase to identify changes in sleep after the intervention. Sleep diaries were also completed for a 2-week period three months after the intervention to determine longer term intervention effects. The authors concluded that sleep quantity and quality improved for all four caregivers and sleep improvement was maintained at 3-month follow-up.

Baseline (A) phases have two purposes; to describe the baseline behaviour and to predict future behaviour if no intervention was applied (Kazdin, 2011). This provides a criterion by which behaviour can be assessed when the intervention is implemented and when the intervention is removed (i.e., based on the information about behaviour from the baseline phase, does behaviour change in a predictable way after the intervention is applied/removed?). As a result, the investigator is able to examine whether the removal of the intervention introduced in the B phase results in the behavioural outcome reverting back to the 'baseline' levels observed in the first A phase. However, in some cases it may not be possible to withdraw the effects of the intervention. Some interventions aim to produce lasting behaviour change (e.g., interventions which provide information about the consequences of performing a behaviour). Therefore, it would not be possible or desirable for the intervention effects to be reversed. None of the ABA studies identified in the review used the final A phase to check for reversal. The purpose of the final A phase in all ABA studies was to check for the long-term maintenance of behaviour change by collecting follow-up measurements weeks or months after the completion of the intervention. Therefore, it is possible that there are other explanations for a sustained change in behaviour aside from the effect of the intervention in these ABA studies. For example, Penica and Williams (2008) discussed the possibility that the improved adherence in a 2-year-old child receiving haemophilia treatment may be related to maturation over time rather than sustained effects of the intervention.

ABAB intervention withdrawal design (and further permutations). Like the baseline phase, the first intervention phase (B) also has two purposes; first to describe the behaviour when the intervention is implemented and second to predict future behaviour if the intervention continued (Kazdin, 2011). An ABAB design has at least two A and two B phases; therefore, it has the opportunity to predict both future non-intervention and future intervention phases. As a result, these designs are often considered as the minimum to demonstrate experimental control (Vohra et al., 2015). Further permutations of ABAB designs involve repeatedly implementing and withdrawing the same intervention. A greater number of AB replications in the design allows a greater number of opportunities to observe whether behaviour changes in the predicted pattern. One study identified in the review used an ABABA intervention withdrawal design (Bernard, Cohen, & Moffett, 2009) and tested the effectiveness of a token economy on increasing and maintaining adherence to specific exercise recommendations in three children with cystic fibrosis. The last A phase involved follow-up measurements at 1- and 3-month post-intervention to check whether the effect of the intervention had been maintained several months after it was finally removed. The authors concluded that two participants increased their exercise levels in response to the token economy. This study makes the assumption that the effect of the intervention can be reversed in the short-term and maintained in the long-term.

Alternating treatment design. An alternating treatment (or condition) design employs similar principles to an ABAB design with the exception that it is used to compare two or more interventions or intervention components rather than compare one intervention against a non-intervention period. This design involves alternating the introduction of different intervention phases (e.g., ABCBCBC) to determine the relative impact of each intervention on outcomes. This can help to determine the best treatment for the individual being studied. An alternating treatment design is also useful for identifying which specific component(s) of the intervention (e.g., behaviour change techniques (BCTs)) are responsible for change or are more effective than others components. Three studies identified in the review used an alternating treatment design to evaluate the impact of interventions and intervention components on treatment adherence (Sonnier, 2002; Vail-Gandolfo, 2009) and physical activity (Cohen, Chelland, Ball, & LeMura, 2002). For example, Sonnier (2002) compared the individual effects of two separate intervention components, self-monitoring and monetary reward, as well as these two intervention components combined, to determine which component(s) was the most effective for increasing treatment adherence in six haemodialysis patients. The authors concluded that only two participants demonstrated increased treatment adherence that could be attributable to either of the treatment variables. To make inferences about whether the order of treatment (or treatment component) was important for eliciting changes in behaviour, Vail-Gandolfo (2009) and Sonnier (2002) allocated half of the participants studied to receive a different order of treatments. Threats to internal validity can be decreased by including a larger number of alternations of treatment conditions. However, only one of the three studies using an alternating treatment design (Cohen et al., 2002) used more than one alternation of the treatment condition tested. In this study, the number of alternations ranged between four and eight alternations across participants.

n-of-1 randomised controlled trial (RCT) design. The n-of-1 RCT design incorporates randomisation to evaluate the effects of one or more interventions (or intervention components) on an individual by randomly allocating different time periods *within individuals* to repeated intervention and control conditions and comparing responses. Randomisation should be incorporated into the n-of-1 design if possible (Tate et al., 2013) because random phase allocation removes potential time-based confounders which are threats to internal validity. Three studies identified in the review used an n-of-1 RCT design to evaluate interventions to change physical activity (Nyman, Goodwin, Kwasnicka, & Callaway, 2015; Sniehotta et al., 2012) and treatment adherence (Lemoncello, 2009). Two studies compared self-monitoring and goal setting, BCTs central to self-regulation theory (Carver & Scheier, 1982), and assessed their effects on walking (Nyman et al., 2015; Sniehotta et al., 2012). Sniehotta et al. (2012) used an n-of-1 RCT 2 × 2 factorial design to test the independent

effects of the two BCTs in 10 adults. The BCTs of goal setting and self-monitoring were delivered randomly to participants over 60 days. Some participants walked more on days when they were instructed to set goals whilst others walked more on days they were instructed to self-monitor behaviour using a pedometer. Nyman et al. (2015) replicated this study in eight older adults and also found variations in individual response to the two BCTs. These studies show that the effectiveness of BCTs on behaviour at the group level may not represent the effectiveness of BCTs for individuals. In another n-of-1 RCT study using an automated TV-based prompt system to increase treatment adherence in patients with dysphasia (Lemoncello, 2009), phases (days) were randomly allocated to intervention (swallowing exercises) or control (typical practice) over time. These n-of-1 RCT studies included between 22 (Lemoncello, 2009) and 60 (Nyman et al., 2015; Sniehotta et al., 2012) phase alternations. Therefore, there is a lower threat to internal validity in these studies compared to other studies identified in the review (i.e., studies using alternating treatment designs).

Strengths of the ABAB design, alternating treatment design and n-of-1 RCT design are that they can be used to promote internal validity and to select the best treatment or treatment components for individuals. This is particularly the case when the design incorporates many alternations between phases and uses randomisation to allocate intervention and control phases. A limitation of these designs is they can only be used in circumstances where the effect of the intervention on behaviour is reversible. In health behaviour research, reversing the effect of the intervention is not always possible or desirable. Even when the effect of an intervention on behaviour can be removed and behaviour will revert to baseline levels, it may not be ethical to remove the intervention. In such cases where reversal is not suitable, multiple baseline designs and changing criterion designs can be used to increase internal validity.

Multiple baseline design. The multiple baseline design involves staggering the introduction of an intervention across participants or across behaviours within the same participant. This method increases internal validity by controlling for history effects in studies where the effects of interventions cannot be easily reversed (Kazdin, 2011). Multiple baseline designs across behaviours within the same individual will measure a target behaviour and a number of control behaviours to identify whether the intervention is effective in changing *only* the target behaviour (whilst the other behaviour(s) remains at baseline levels). Multiple baseline designs across participants are either 'concurrent' or 'non-concurrent'. Concurrent designs are when all participants start the study at the same time, with different baseline phase lengths so that the implementation of the intervention remains staggered. 'Non-concurrent' designs involve participants starting the study at different time points and this may be preferred because it allows more flexibility in recruitment.

Eight studies identified in the review used a multiple baseline design to evaluate interventions; six of the studies used a multiple baseline design across individuals (Gorczyński, Morrow, & Irwin, 2008; Pazuano-Slamm, 2005; Smith, 2012b; Ward, Brinkman, Slifer, & Paranjape, 2010; Yngman-Uhlin, Fernstrom, Borjeson, & Edell-Gustafsson, 2012) or individual units (i.e., care centre sites; De Marco, Zeisel, & Odom, 2015) and two of the studies used a multiple baseline design across behaviours (Lane-Brown & Tate, 2010; Romero, 2010). The multiple baseline studies tested interventions to change physical activity, sleep, treatment adherence, smoking and drug use.

Threats to internal validity are decreased with a larger number of comparisons (i.e., more individuals or control behaviours). Four of the five studies with a multiple baseline design across individuals compared the intervention across more than two people. Of the two studies with a multiple baseline design across behaviours, Romero (2010) measured one control behaviour and Lane-Brown and Tate (2010) measured two control behaviours in addition to the target behaviour. To increase the rigour of the multiple baseline design, intervention phase commencement can be randomised (Tate et al., 2013). However, randomisation may not always be possible. Two of the studies reported that randomisation was not possible due to ethical concerns of withholding treatment (Lane-Brown & Tate, 2010; Romero, 2010). By definition, using multiple baseline designs enables multiple opportunities to determine whether the intervention is effective. However, some individuals will have to wait longer than others for treatment which may be unethical or harmful to the individual.

Changing criterion design. The changing criterion design also increases experimental control. It involves phase changes when the outcome variable meets certain pre-selected criteria which can be raised or lowered as the study progresses. This design is often used when gradual changes in behaviour are desirable (Kazdin, 2011). If the intervention is effective, the outcome variable will demonstrate gradual changes corresponding to the criterion selected. One study identified in the review used a changing criterion design (Cervantes & Porretta, 2013). The study evaluated the impact of an after-school programme on physical activity among adolescents with visual impairments using a changing physical activity goal, the achievement of which allowed movement into the next intervention session.

General methodological considerations

Measuring behaviour over time

A crucial aim of behavioural assessment in n-of-1 studies is to obtain a representative (i.e., valid and reliable) picture of the behaviour. There are a number of considerations to take into account in relation to the length of phases, the number/frequency of measurements within phases and the method of measurement.

The degree of variability in the behaviour determines the length of phases and number of measurements required. Highly variable outcomes require measurement for a longer duration or more frequently (Bolger & Laurenceau, 2013), in comparison to outcomes which are more stable over time. The length of phases and number of measurements selected is partly determined by the nature of the behaviour (e.g., alcohol consumption may vary more day-to-day than tooth brushing) and partly determined by individuals (e.g., some people's walking behaviour is highly stable whereas for others it is highly variable). There are instances in which fewer measurements of the behaviour may be appropriate. For example, one study identified in the review measured drug use which occurred infrequently (Norberg et al., 2014); therefore, it was not necessary for the participants to complete measurements on a daily basis. In some cases, it may be unnecessary or unethical to measure the behaviour at all. For example, if an individual's current behaviour is dangerous, it would be unethical to withhold treatment to measure the behaviour over a 'baseline' period. Indeed, one study identified in the review reported ethical concerns about delaying intervention for a child with haemophilia who would not comply with medical treatment (Penica & Williams, 2008). The child refused to comply with any part of the treatment so it was not necessary to prospectively measure compliance behaviour during a baseline (A) phase.

It is also important to consider the presence of trends (i.e., steady increase or decrease during the phase; weekly trends), carryover effects and participant reactivity. Trends in the data can interfere with data interpretation. If a trend is present in the baseline data in the same direction as the direction of the change the intervention aims to achieve, it becomes difficult to identify whether the intervention changed the outcome behaviour, or if it was changing on its own anyway (Kazdin, 2011). A carryover effect is when the effect of the intervention in one phase carries over into the next phase and a key challenge is forming hypotheses about carryover effects that enable the investigator to adequately design the length of intervention and control phases (McDonald et al., 2016a). Participant reactivity refers to when participants change their behaviour due to the awareness of being monitored. It often occurs at the start of a study and can lead to inaccurate baseline data against which to compare intervention phase data. One study identified in the review (Gray et al., 2011) addressed participant reactivity by permitting participants to stop the baseline phase and start the intervention phase only once their medication adherence behaviour was more typical of the medication adherence rates recorded in their medical notes. In another study (Cervantes & Porretta, 2013), participants were permitted to move to the next phase only when their data was 'stable' (i.e., not showing a trend) within the last three measurements. If the investigator identified participant reactivity in some of the data within the baseline phase, these data may be excluded from the analyses (Nyman et al., 2015).

It is often necessary to include several measurements of behaviour over a substantial period of time to account for variability, trends, carryover effects and participant reactivity. The variability in the data can be unique to individuals. Measurements may need to be tailored to each individual in order to detect variability. Running a 'pilot' phase before the start of the study helps to inform the length of phases and number of measurements required.

The method of measurement is a key consideration for obtaining a representative picture of behaviour. Objective methods of assessing behaviour should be used where available, as they are often more valid and reliable than subjective methods, which are prone to a number of errors and biases. Advances in technology have resulted in a wider range of objective measurement methods, including mobile phones, sensors and psychophysiological measures (e.g., carbon monoxide monitors to verify smoking cessation). However, some behaviours and predictors (e.g., cognitions or affect) can be measured only by self-report. Combining self-report with sampling techniques such as EMA to enable real-time sampling in the participant's natural environment can reduce retrospective reporting biases (for a detailed discussion on EMA methods see Shiffman, Stone, & Hufford, 2008). No studies identified in the review employed EMA methods but a number of n-of-1 studies using EMA are being currently being conducted (McDonald, Vieira, O'Brien, White, & Sniehotta, 2016b; Newham, Presseau, Araujo-Soares, & Sniehotta, 2015). A greater number of self-report measurements may increase measurement burden. However, this may be mitigated by their personal relevance to the individual.

Analysing n-of-1 research

To date, the vast majority of published n-of-1 studies have used visual ($n = 21$) or descriptive analysis ($n = 7$) (i.e., narrative description of change in average score across phases) to evaluate n-of-1 data. Only a small proportion have fully exploited the possibilities of n-of-1 studies by utilising statistical methods ($n = 11$).

Visual analysis. Visual analysis involves visually inspecting graphed data. A number of important features are considered when using visual analysis to identify the presence of a functional relationship between predictor and outcome variables including changes in level (i.e., average), slope (i.e., trend), variability, immediacy of effect, overlap between phases and consistency of data in similar phases (Chen, Peng, & Chen, 2015). Studies identified in the review using visual analysis examined changes in level across phases (Bernard et al., 2009; Cervantes & Porretta, 2013; Cohen et al., 2002; Gorczynski et al., 2008; Gorski et al., 2005; Gray et al., 2011; Hill et al., 2015; Lee et al., 2014; MacPherson et al., 2016; McCrae et al., 2005; Norberg et al., 2014; Payne et al., 2013; Penica & Williams, 2008; Piven & Duran, 2014; Romero, 2010; Seveck et al., 2005; Smith, 2012b; Sonnier, 2002; Soroudi et al., 2008; Ward et al., 2010; Yngman-Uhlin et al., 2012). Five studies also examined changes in slope across phases (Gorczynski et al., 2008; Gray et al., 2011; Payne et al., 2013; Smith, 2012b; Vail-Gandolfo, 2009), two studies examined the latency of the intervention effect (Gorczynski et al., 2008; Smith, 2012b) and one examined overlap in data points across different phases (Gorczynski et al., 2008).

Statistical analyses. A recent meta-analysis demonstrated low inter-rater agreement when interpreting graphed n-of-1 data (Ninci, Vannest, Willson, & Zhang, 2015). When applicable, statistical analyses are more reliable for determining whether there have been changes in trends and variability between phases and for detecting gradual changes and small effects (Kazdin, 2011). Statistical techniques used in the studies identified in the review included two-sample paired and unpaired *t*-tests (Pauzano-Slamm, 2005), Cohen's *d* (Cortina et al., 2013; Lemoncello, 2009), classical test theory (Vail-Gandolfo, 2009), tau-*u* index (De Marco et al., 2015), cross-correlational analysis (Hobbs et al., 2013; Quinn et al., 2013), regression-based analysis (Cortina et al., 2013; Lane-Brown & Tate, 2010; Nyman et al., 2015; O'Brien et al., 2016; Quinn et al., 2013; Sniehotta et al., 2012) and simulation modelling analysis (O'Brien et al., 2016) (further details are provided in Supplemental Table 2).

Prior to carrying out statistical analyses, statistical power should be considered. In n-of-1 research, statistical power is a function of the number of data points collected from one individual, and is highly dependent on intra-individual variability and autocorrelation. Autocorrelated data mean that the

present observation can be explained by previous observations (e.g., pain today is partially predicted by pain yesterday). In other words, the observations are not independent from each other and each observation adds less information as it partially reflects previous observations. Autocorrelation is common when repeated measurements are taken from the same individual. Applying conventional statistical methods (e.g., *t*-test) to autocorrelated data may inflate type-1 error probability. There is no universal guideline describing the optimal or minimum number of data points required for n-of-1 research, although a minimum of 50 observations per phase has been proposed for some statistical techniques which may be a useful starting point (Tabachnick & Fidell, 2007). However, statistical power is specific to the individual since it is related to the effect size, standard deviation and number of observations. Few studies identified in the review justified the statistical power of their analyses or accounted for autocorrelation in the analysis.

Statistical analyses of n-of-1 data should include accurate modelling of the outcome variable while accounting for its distribution, time-related trend and error structures (e.g., autocorrelation) as well as reporting effect sizes that are easily understood and usable for clinical decision-making. A substantial number of statistical approaches have been documented but no clear consensus exists on which method is most appropriate for which kind of design and data. One study identified in the review used conventional parametric tests to compare changes in behaviours within individuals but did not test for autocorrelation (Pauzano-Slamm, 2005). In some statistical approaches, autocorrelation is identified and adjusted for, prior to analyses (a procedure called 'pre-whitening'). Three studies used a pre-whitening procedure to remove autocorrelation from predictors and outcomes prior to analysis (Hobbs et al., 2013; Nyman et al., 2015; Quinn et al., 2013) which involved cross-correlational analysis and multiple linear regression.

The features of the data (e.g., type of outcome, time-trend, autocorrelation and seasonality patterns) are likely to be unique to the individual and the data may need to be transformed using unique protocols (e.g., pre-whitening) before statistical analysis. As a result, it may be difficult to have a full a priori plan for the analysis of n-of-1 data, rather a number of options depending on the design used. Missing data are also a challenge for using statistical techniques. Methods for imputing missing data may be required prior to analysis. Of the studies identified in the review, few reported methods for imputing missing data. There are various approaches to data imputation such as simple and multiple imputation which have been well described elsewhere (Schafer & Graham, 2002). Three of the studies identified in the review (Hobbs et al., 2013; Nyman et al., 2015; Quinn et al., 2013) reported the use of multiple imputation methods for missing data.

In n-of-1 research, the generalisability of findings (over time within individuals or to other individuals) relies on replication of n-of-1 studies within and between individuals. Recommendations for the minimum number of replications needed before an intervention effect can be generalised across individuals have been published in other fields (Kratochwill et al., 2013). However, the picture is more complex than these would suggest as it depends on whether replication studies produce similar results, the observed effect size, statistical power and *p* value. There has been progress towards optimising multilevel modelling and meta-analytic procedures for combining n-of-1 data from different individuals (Manolov, Gast, Perdices, & Evans, 2014). Most articles identified in the review reported on a series of n-of-1 studies with different participants (median 4 participants, range 2–14). Only two studies used statistical procedures to combine the data from individual participants, in addition to analyses conducted at the individual level (Nyman et al., 2015; Sniehotta et al., 2012).

Discussion

Main findings

This article provided an overview of the use of n-of-1 methods to study or change health-related behaviour using illustrative examples identified using a systematic review. n-of-1 methods have

been used across a wide variety of population groups, of various ages, and health-related behaviours, including treatment adherence, physical activity, recreational drug use, smoking, sleep, eating behaviour and alcohol use. However, to date, investigators have not utilised the more sophisticated options of n-of-1 methods in health behaviour research. There is scope to apply n-of-1 methods to health behaviour research more widely, particularly using some of the more sophisticated and rigorous n-of-1 designs (e.g., n-of-1 RCTs) and newer methods of statistical analysis. The articles identified by the systematic review highlighted a number of opportunities and challenges for future n-of-1 research.

Testing theories about behaviour

Most theories about behaviour and behaviour change describe the behaviour of individuals, yet they have been tested using between-participant study designs (Johnston & Johnston, 2013). These designs cannot establish whether the assumptions of a theory are supported at the individual level (Molenaar & Campbell, 2009). n-of-1 designs can be used to test theory within individuals. In n-of-1 studies, TPB variables do not predict the physical activity behaviour for all individuals and, in some cases, the TPB constructs predict behaviour in the opposite direction to that predicted by the theory (Hobbs et al., 2013; Quinn et al., 2013). These findings are in contrast with the large body of evidence from group-based research (McEachan, Conner, Taylor, & Lawton, 2011).

n-of-1 interventional designs can also be used to test theory at the individual level by manipulating the causal determinants of behaviour (Medical Research Council, 2008), such as those proposed by current theories of behaviour, to identify their predictive validity across individuals and behaviours. However, no studies of this kind were identified in the review. Furthermore, the temporal ordering of cognitions and behaviour provides a precondition for an indication of the likely causal relationships between cognitions and behaviour, information that is of use for intervention design. If causality can be demonstrated by experimental methods, an understanding of the time course of the causal relationship between cognitions and behaviour would enable the timing of intervention delivery to be optimised.

Evaluating interventions

Testing interventions using traditional between-participant designs provides information about the average effect of interventions at the group level but little about responses to interventions at the individual level (Davidson, Peacock, Kronish, & Edmondson, 2014). Group level results have been shown to misrepresent individual response (Nyman et al., 2015; Sniehotta et al., 2012). Knowledge about individual response can lead to improvement in intervention tailoring and optimisation, that is, the selection of the best intervention for an individual based on their own empirical data. The unique determinants of behaviour for an individual can be identified during an observational phase and this information can be used to develop an individualised intervention plan targeting these determinants (O'Brien et al., 2016).

n-of-1 methods can be applied to the study of processes and outcomes (Barlow, Nock, & Hersen, 2009). Obtaining information about the process of behaviour change (e.g., mechanisms, timing and nature of behaviour change) within individuals is a particular strength of n-of-1 methods as this information can be obscured in between-participant studies. n-of-1 methods can be used to test intervention components individually or in combination with others and to test different combinations, sequences and doses of intervention components to identify the active ingredients required for long-term changes in behaviour (McDonald et al., 2016a). A systematic search of the literature identified few n-of-1 RCTs, yet these designs can make a considerable contribution to the debate around methods which are effective in changing behaviour (e.g., BCTs). Sequential repetition of intervention and control periods are characteristic for n-of-1 studies and this is a

particularly attractive feature for behavioural science when studying maintenance and dose–response relationships (e.g., how many times do I need to monitor my behaviour to develop a habit?).

Personalising health psychology and behavioural medicine

Opportunities to advance behavioural science are facilitated by rapid technological developments which allow investigators to capture large amounts of data and measure behaviours and predictors unobtrusively over time within individuals (Dallery, Cassidy, & Raiff, 2013). n-of-1 study designs are flexible, enabling the design to be personalised to the interests and requirements of the individual. Results from an n-of-1 study can be shared and discussed with the participant (Kwasnicka, Domrowski, White, & Sniehotta, 2015). This may encourage individuals to play an active role in their health and can contribute to a process of knowledge co-creation. Practitioners can use n-of-1 methods as a decision-making tool with patients, which can result in the selection of better treatment (Joy et al., 2014). Thus, n-of-1 methods are consistent with the movement towards person-centred and personalised medicine (Lillie et al., 2011). n-of-1 methods are recognised, alongside systematic reviews of RCTs, as providing the highest level of evidence for making clinical decisions for individuals (Howick et al., 2011) and are particularly useful for studying small patient populations and rare conditions (Lillie et al., 2011), demonstrating their exciting potential in research and in practice.

Strengths and limitations

This is the first comprehensive overview that describes using n-of-1 methods to study or change health behaviours. The article provided illustrative examples from a systematic review that used rigorous methods to identify the breadth of types and uses of n-of-1 studies in health behaviour research. However, some limitations are acknowledged. Numerous terms have been used to describe n-of-1 methods in the literature and some specific terms (e.g., ‘multiple baseline’) were not explicitly included in the search strategy. These specific terms are labels used to describe sub-types of n-of-1 studies; therefore, studies using these terms will also use broader terms (e.g., single-case study, n-of-1) which were included in the search. Given that the review was conducted to illustrate cases, rather than to estimate parameters, conclusions are likely to be robust. The systematic review included only published articles and dissertations so there may be a risk of publication bias. Only articles published after January 2000 were eligible for inclusion so there may also be a risk of excluding relevant studies published before this date. However, many of the methods of assessing behaviour require technology, which has rapidly advanced during this period (Dallery et al., 2015).

Implications for future research and practice

The systematic review of n-of-1 studies in health behaviour research highlighted a number of unmet challenges and unanswered questions which should be considered in future research. While specifying an n-of-1 protocol, investigators must make trade-offs between several factors including feasibility, measurement characteristics, scientific rigour and the specificity of the n-of-1 design. Statistical power represents a significant challenge. For example, a priori power calculations may not be possible because statistical power (i.e., number of data points) is likely to differ between individuals. This has implications for informed consent procedures due to uncertainty about how long the individual needs to participate for.

Statistical analysis is needed in many n-of-1 studies using a high resolution sampling of outcomes. However, few articles used statistical approaches for data analysis, and in those cases there was a tendency to use simpler and, arguably, less appropriate statistical methods. Although several statistical approaches are available for determining treatment effect sizes in n-of-1 studies,

little consensus exists about which approach to use and in which circumstances (Vohra et al., 2015). Currently available statistical techniques rely heavily on features of the data (i.e., variability, linearity, normality, autocorrelation) which are likely to differ between individuals. This may suggest that unique protocols are required for each participant which would have an impact on the capacity to statistically combine cases in meta-analyses. Recently, a number of new statistical approaches have been developed (e.g., Lin et al., 2016), which has opened avenues to more sophisticated n-of-1 statistical analyses, but further validation of their performance is needed. Furthermore, their usability will depend on accessibility and user-friendliness. Future research should focus on the development and use of statistical techniques which are adaptable to different types of outcomes and capable of dealing with different challenges inherent to n-of-1 data modelling. Such methods should allow the identification of predictors of response, the description of adaptive changes over time and prediction of future behaviour given prior history as well as explicitly investigating carryover effects.

Although the acceptability of n-of-1 pharmacological studies to patients has been explicitly considered in the medical field (Nikles, Clavarino, & Del Mar, 2005), future research should explore the acceptability of n-of-1 studies to study or change health behaviour. Future research should also identify how n-of-1 methods can be used to study and change other health-related outcomes such as symptoms including pain and fatigue.

n-of-1 methods are historically and predominantly a practice-based method that practitioners have brought to scientists rather than vice versa. Therefore, it can be assumed that in various clinical settings practitioners use n-of-1 methods. This overview may help practitioners to optimise the rigour of the conclusions they reach about individuals.

Conclusions

n-of-1 methods are recognised as a viable and versatile research method in various disciplines, but have so far been under-recognised and under-used in health psychology and behavioural medicine. Although the use of n-of-1 designs in health behaviour research is still in its infancy, this article highlights a range of relevant issues for using n-of-1 methods in future research and practice. n-of-1 methods are an important addition to the repertoire of study designs and can provide the tools needed to advance theory and personalise health behaviour interventions to individuals. n-of-1 methods provide opportunities to answer some of the key questions in health behaviour research.

Acknowledgements

MW is a member of the Centre for Diet and Activity Research (CEDAR). Both Fuse and CEDAR are UK Clinical Research Collaboration (UKCRC) Public Health Research Centres of Excellence. Funding from the British Heart Foundation, Cancer Research UK, Economic and Social Research Council, Medical Research Council, the National Institute for Health Research, and the Wellcome Trust under the auspices of the UKCRC, is gratefully acknowledged.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was part of the LiveWell programme [grant number G0900686]. LiveWell was supported by the Lifelong Health and Wellbeing initiative (LLHW). The LLHW initiative is a funding collaboration between the UK Research Councils and Health Departments. The funding partners are Biotechnology and Biological Sciences Research Council, Engineering and Physical Sciences Research Council, Economic and Social Research Council, Medical Research Council, Chief Scientist Office of the Scottish Government Health Directorates, National Institute for Health Research/The Department of

Health, The Health and Social Care Research & Development of the Public Health Agency (Northern Ireland), and Wales Office of Research and Development for Health and Social Care, Welsh Assembly Government. FFS is funded by Fuse, the UK Clinical Research Collaboration Centre for Translational Research in Public Health.

References

- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50(2), 179–211. doi:10.1016/0749-5978(91)90020-T
- Araujo, A., Julious, S., & Senn, S. (2016). Understanding variation in sets of n-of-1 trials. *Plos One*, 11(12). doi:10.1371/journal.pone.0167167
- Banducci, A. N., Long, K. E., & MacPherson, L. (2015). A case series of a behavioral activation-enhanced smoking cessation program for inpatient substance users with elevated depressive symptoms. *Clinical Case Studies*, 14(1), 61–77. doi:10.1177/1534650114538699
- Barlow, D. H., Nock, M., & Hersen, M. (2009). *Single case experimental designs: Strategies for studying behavior for change* (3rd ed.). Boston, MA: Pearson/Allyn and Bacon.
- Bernard, R. S., Cohen, L. L., & Moffett, K. (2009). A token economy for exercise adherence in pediatric cystic fibrosis: A single-subject analysis. *Journal of Pediatric Psychology*, 34(4), 354–365. doi:10.1093/jpepsy/jsn101
- Bolger, N., & Laurenceau, J.-P. (2013). *Intensive longitudinal methods: An introduction to diary and experience sampling research*. New York, NY: Guilford Press.
- Carver, C. S., & Scheier, M. F. (1982). Control theory: A useful conceptual framework for personality-social, clinical, and health psychology. *Psychological Bulletin*, 92(1), 111–135. doi:10.1037//0033-2909.92.1.111
- Casey, A. F., Mackay-Lyons, M., Connolly, E. M., Jennings, C., & Rasmussen, R. (2014). A comprehensive exercise program for a young adult male with Down's syndrome who experienced a stroke. *Disability and Rehabilitation*, 36(17), 1402–1408. doi:10.3109/09638288.2013.833304
- Cervantes, C. M., & Porretta, D. L. (2013). Impact of after school programming on physical activity among adolescents with visual impairments. *Adapted Physical Activity Quarterly*, 30(2), 127–146. doi:10.1123/apaq.30.2.127
- Chen, L. T., Peng, C. Y., & Chen, M. E. (2015). Computing tools for implementing standards for single-case designs. *Behavior Modification*, 39(6), 835–869. doi:10.1177/0145445515603706
- Cohen, S., Chelland, Ball, K., & LeMura, L. (2002). Effects of fixed ratio schedules of reinforcement on exercise by college students. *Perceptual and Motor Skills*, 94(3 Pt 2), 1177–1186. doi:10.2466/pms.2002.94.3c.1177
- Cortina, S., Somers, M., Rohan, J. M., & Drotar, D. (2013). Clinical effectiveness of comprehensive psychological intervention for nonadherence to medical treatment: A case series. *Journal of Pediatric Psychology*, 38(6), 649–663. doi:10.1093/jpepsy/jss175
- Dallery, J., Cassidy, R. N., & Raiff, B. R. (2013). Single-case experimental designs to evaluate novel technology-based health interventions. *Journal of Medical Internet Research*, 15(2), doi:10.2196/jmir.2227
- Dallery, J., Kurti, A., & Erb, P. (2015). A new frontier: Integrating behavioral and digital technology to promote health behavior. *Behavior Analyst*, 38(1), 19–49. doi:10.1007/s40614-014-0017-y
- Daughters, S. B., Magidson, J. F., Schuster, R. M., & Safren, S. A. (2010). Act Healthy: A combined cognitive-behavioral depression and medication adherence treatment for HIV-infected substance users. *Cognitive and Behavioral Practice*, 17(3), 309–321. Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3122920/>
- Davidson, K. W., Peacock, J., Kronish, I. M., & Edmondson, D. (2014). Personalizing behavioral interventions through single-patient (N-of-1) trials. *Social and Personality Psychology Compass*, 8(8), 408–421. doi:10.1111/spc3.12121
- De Marco, A. C., Zeisel, S., & Odom, S. L. (2015). An evaluation of a program to increase physical activity for young children in child care. *Early Education and Development*, 26(1), 1–21. doi:10.1080/10409289.2014.932237
- Gabler, N. B., Duan, N., Vohra, S., & Kravitz, R. L. (2011). N-of-1 trials in the medical literature: A systematic review. *Medical Care*, 49(8), 761–768. doi:10.1097/MLR.0b013e318215d90d
- Gonzalez, J. S., McCarl, L. A., Wexler, D. J., Cagliero, E., Delahanty, L., Soper, T. D., ... Safren, S. A. (2010). Cognitive-behavioral therapy for adherence and depression (CBT-AD) in type 2 diabetes. *Journal of Cognitive Psychotherapy*, 24(4), 329–343. doi:10.1891/0889-8391.24.4.329
- Gorczynski, P., Morrow, D., & Irwin, J. D. (2008). The impact of co-active coaching on physically inactive 12 to 14 year olds in Ontario. *International Journal of Evidence Based Coaching and Mentoring*, 6(2), 13–26. Retrieved from <http://jebcm.brookes.ac.uk/documents/vol06issue2-paper-02.pdf>
- Gorski, J. A., Slifer, K. J., Townsend, V., Kelly-Suttka, J., & Amari, A. (2005). Behavioural treatment of non-compliance in adolescents with newly acquired spinal cord injuries. *Pediatric Rehabilitation*, 8(3), 187–198. doi:10.1080/13638490400021438
- Gray, W. N., Janicke, D. M., Fennell, E. B., Driscoll, D. C., & Lawrence, R. M. (2011). Piloting behavioral family systems therapy to improve adherence among adolescents with HIV: A case series intervention study. *Journal of Health Psychology*, 16(5), 828–842. doi:10.1177/1359105310394230
- Hill, M. L., Masuda, A., Moore, M., & Twohig, M. P. (2015). Acceptance and commitment therapy for individuals with problematic emotional eating: A case-series study. *Clinical Case Studies*, 14(2), 141–154. doi:10.1177/1534650114547429

- Hobbs, N., Dixon, D., Johnston, M., & Howie, K. (2013). Can the theory of planned behaviour predict the physical activity behaviour of individuals? *Psychology and Health*, 28(3), 234–249. doi:10.1080/08870446.2012.716838
- Howick, J., Chalmers, I., Glasziou, P., Greenhaigh, T., Heneghan, C., Liberati, A., & Thornton, H. (2011). *The 2011 oxford CEBM evidence table*. Oxford Centre for evidence-based medicine. Retrieved from <http://www.cebm.net/wp-content/uploads/2014/06/CEBM-Levels-of-Evidence-2.1.pdf>
- Hox, J. J., Moerbeek, M., & van de Schoot, R. (2010). *Multilevel analysis: Techniques and applications* (2nd ed.). New York: Routledge.
- Johnston, D. W., & Johnston, M. (2013). Useful theories should apply to individuals. *British Journal of Health Psychology*, 18(3), 469–473. doi:10.1111/bjhp.12049
- Johnston, M., & Dixon, D. (2014). Developing an integrated biomedical and behavioural theory of functioning and disability: Adding models of behaviour to the ICF framework. *Health Psychology Review*, 8(4), 381–403. doi:10.1080/17437199.2013.855592
- Joy, T. R., Monjed, A., Zou, G. Y., Hegele, R. A., McDonald, C. G., & Mahon, J. L. (2014). N-of-1 (single-patient) trials for statin-related myalgia. *Annals of Internal Medicine*, 160(5), 301–310. doi:10.7326/M13-1921
- Lane-Brown, A., & Tate, R. (2010). Evaluation of an intervention for apathy after traumatic brain injury: A multiple-baseline, single-case experimental design. *Journal of Head Trauma Rehabilitation*, 25(6), 459–469. doi:10.1097/HTR.0b013e3181d98e1d
- Lee, D. C., Budney, A. J., Brunette, M. F., Hughes, J. R., Etter, J. F., & Stanger, C. (2014). Treatment models for targeting tobacco use during treatment for cannabis use disorder: Case series. *Addictive Behaviors*, 39(8), 1224–1230. doi:10.1016/j.addbeh.2014.04.010
- Lemoncello, R. R. (2009). *A within-subjects experimental evaluation of the television assisted prompting (TAP) system to maximize completion of home-delivered swallow strengthening exercises among individuals with co-occurring acquired swallowing and cognitive impairments* (Unpublished PhD dissertation). University of Oregon, Oregon.
- Lowe, S. S., Watanabe, S. M., Baracos, V. E., & Courneya, K. S. (2013). Home-based functional walking program for advanced cancer patients receiving palliative care: A case series. *BMC Palliative Care*, 12(1). doi:10.1186/1472-684X-12-22
- Kazdin, A. E. (2011). *Single-case research designs: Methods for clinical and applied settings* (2nd ed.). New York: Oxford University Press.
- Kratochwill, T. R., Hitchcock, J. H., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2013). Single-case intervention research design standards. *Remedial and Special Education*, 34(1), 26–38. doi:10.1177/0741932512452794
- Kwasnicka, D., Dombrowski, S. U., White, M., & Sniehotta, F. F. (2015). Data-prompted interviews: Using individual ecological data to stimulate narratives and explore meanings. *Health Psychology*, 34(12), 1191–1194. doi:10.1037/hea0000234
- Lillie, E. O., Patay, B., Diamant, J., Issell, B., Topol, E. J., & Schork, N. J. (2011). The n-of-1 clinical trial: The ultimate strategy for individualizing medicine? *Personalized Medicine*, 8(2), 161–173. doi:10.2217/pme.11.7
- Lin, S. X., Morrison, L., Smith, P. W., Hargood, C., Weal, M., & Yardley, L. (2016). Properties of bootstrap tests for N-of-1 studies. *British Journal of Mathematical and Statistical Psychology*. doi:10.1111/bmsp.12071
- MacPherson, L., Collado, A., Ninnemann, A., & Hoffman, E. (in press). Development of a behavioral activation-based intervention for cigarette-smoking young adults. *Cognitive and Behavioral Practice*. doi:10.1016/j.cbpra.2016.03.004
- Manolov, R., Gast, D. L., Perdices, M., & Evans, J. J. (2014). Single-case experimental designs: Reflections on conduct and analysis. *Neuropsychological Rehabilitation*, 24(3), 634–660. doi:10.1080/09602011.2014.903199
- McCrae, C. S., Tierney, C. G., & McNamara, J. P. H. (2005). Behavioral intervention for insomnia: Future directions for non-traditional caregivers at various stages of care. *Clinical Gerontologist*, 29(2), 95–114. doi:10.1300/J018v29n02_10
- McDonald, S., Araujo-Soares, V., & Sniehotta, F. F. (2016a). N-of-1 randomised controlled trials in health psychology and behavioural medicine. *Psychology & Health*, 31(3), 331–333. doi:10.1080/08870446.2016.1145221
- McDonald, S., & Davidson, K. W. (2016). Using N-of-1 methods to study or change health-related behaviour and outcomes: A symposium summary. *European Health Psychologist*, 18, 38–42. Retrieved from <http://www.ehps.net/ehp/index.php/contents/article/view/1415/>
- McDonald, S., Quinn, F., Hobbs, N., White, M., & Sniehotta, F. F. (2014). *A systematic review of n-of-1 observational and experimental studies with health-related behavioural outcomes*. Retrieved from http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42014007258
- McDonald, S., Vieira, R., O'Brien, N., White, M., & Sniehotta, F. F. (2016b). Does physical activity and sedentary behavior change during the retirement transition? Findings from a series of novel n-of-1 natural experiments. *International Journal of Behavioral Medicine*, 23, S261–S261. doi:10.1007/s12529-016-9586-3
- McEachan, R. R. C., Conner, M., Taylor, N. J., & Lawton, R. J. (2011). Prospective prediction of health-related behaviours with the theory of planned behaviour: A meta-analysis. *Health Psychology Review*, 5(2), 97–144. doi:10.1080/17437199.2010.521684
- Medical Research Council. (2008). *Developing and evaluating complex interventions: New guidance*. Retrieved from <https://www.mrc.ac.uk/documents/pdf/complex-interventions-guidance/>
- Moeller, J. D., Dattilo, J., & Rusch, F. (2015). Applying quality indicators to single-case research designs used in special education: A systematic review. *Psychology in the Schools*, 52(2), 139–153. doi:10.1002/pits.21801
- Molenaar, P. C. M., & Campbell, C. G. (2009). The new person-specific paradigm in psychology. *Current Directions in Psychological Science*, 18(2), 112–117. doi:10.1111/j.1467-8721.2009.01619.x

- Newham, J., Presseau, J., Araujo-Soares, V., & Sniehotta, F. F. (2015). Differential responses to interventions targeting physical activity compared to sedentary time: A randomized n-of-1 design. *European Health Psychologist*, 17, 520. Retrieved from <http://www.ehps.net/ehp/index.php/contents/article/view/890>
- Nikles, C. J., Clavarino, A. M., & Del Mar, C. B. (2005). Using n-of-1 trials as a clinical tool to improve prescribing. *British Journal of General Practice*, 55(512), 175–180. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/15808031>
- Ninci, J., Vannest, K. J., Willson, V., & Zhang, N. (2015). Interrater agreement between visual analysts of single-case data: A meta-analysis. *Behavior Modification*, 39(4), 510–541. doi:10.1177/0145445515581327
- Norberg, M. M., Perry, U., Mackenzie, J., & Copeland, J. (2014). Met plus CBT for ecstasy use when clients are depressed: A case series. *Cognitive and Behavioral Practice*, 21(1), 55–63. doi:10.1016/j.cbpra.2013.06.002
- Norcross, J. C., & Wampold, B. E. (2011). What works for whom: Tailoring psychotherapy to the person. *Journal of Clinical Psychology*, 67(2), 127–132. doi:10.1002/jclp.20764
- Nyman, S. R., Goodwin, K., Kwasnicka, D., & Callaway, A. (2015). Increasing walking among older people: A test of behaviour change techniques using factorial randomised N-of-1 trials. *Psychology & Health*, 1–18. doi:10.1080/08870446.2015.1088014
- O'Brien, N., Philpott-Morgan, S., & Dixon, D. (2016). Using impairment and cognitions to predict walking in osteoarthritis: A series of n-of-1 studies with an individually tailored, data-driven intervention. *British Journal of Health Psychology*, 21(1), 52–70. doi:10.1111/bjhp.12153
- Pauzano-Slamm, N. M. (2005). *Mindfulness meditation for chronic fatigue syndrome: A controlled trial* (Unpublished PhD dissertation). Hofstra University, New York.
- Payne, M. E., Eaton, C. K., Mee, L. L., & Blount, R. L. (2013). Promoting medication adherence and regimen responsibility in two adolescents on hemodialysis for end stage renal disease: A case study. *Clinical Case Studies*, 12(2), 95–110. doi:10.1177/1534650112467079
- Penica, S., & Williams, K. E. (2008). The use of a psychological intervention to increase adherence during factor administration in a child with haemophilia. *Haemophilia*, 14(5), 939–944. doi:10.1111/j.1365-2516.2008.01789.x
- Perdices, M., & Tate, R. L. (2009). Single-subject designs as a tool for evidence-based clinical practice: Are they unrecognised and undervalued? *Neuropsychological Rehabilitation*, 19(6), 904–927. doi:10.1080/09602010903040691
- Piven, E., & Duran, R. (2014). Reduction of non-adherent behaviour in a Mexican-American adolescent with type 2 diabetes. *Occupational Therapy International*, 21(1), 42–51. doi:10.1002/oti.1363
- Punja, S., Bukutu, C., Shamseer, L., Sampson, M., Hartling, L., Urichuk, L., & Vohra, S. (2016). The design, analysis and meta-analysis of N-of-1 trials: A tapestry of heterogeneity. *Journal of Clinical Epidemiology*. doi:10.1016/j.jclinepi.2016.03.023
- Quinn, F., Johnston, M., & Johnston, D. W. (2013). Testing an integrated behavioural and biomedical model of disability in n-of-1 studies with chronic pain. *Psychology & Health*, 28(12), 1391–1406. doi:10.1080/08870446.2013.814773
- Romero, V. I. (2010). *Parental substance abuse and child neglect: A controlled trial of a developed treatment manual* (Unpublished PhD dissertation). University of Nevada, Las Vegas.
- Sather, B. C., Forbes, J. J., Starck, D. J., & Rovers, J. P. (2007). Effect of a personal automated dose-dispensing system on adherence: A case series. *Journal of the American Pharmacists Association*, 47(1), 82–85. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/17338479>
- Schaffer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, 7(2), 147–177. doi:10.1037/1082-989X.7.2.147
- Scholes, S., & Mindell, J. (2013). *Health survey for England 2012. Chapter 2: Physical activity in adults*. London: The Health and Social Care Information Centre.
- Sevick, M. A., Piraino, B., Sereika, S., Starrett, T., Bender, C., Bernardini, J., ... Burke, L. E. (2005). A preliminary study of PDA-based dietary self-monitoring in hemodialysis patients. *Journal of Renal Nutrition*, 15(3), 304–311. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/16007560>
- Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment. *Annual Review of Clinical Psychology*, 4, 1–32. doi:10.1146/annurev.clinpsy.3.022806.091415
- Smith, J. D. (2012a). Single-case experimental designs: A systematic review of published research and current standards. *Psychological Methods*, 17(4), 510–550. doi:10.1037/a0029312
- Smith, R. C. (2012b). *Test of a treatment for PTSD and nicotine dependence using intensive, single-case designs* (Unpublished PhD dissertation). University of Arkansas, Fayetteville.
- Sniehotta, F. F., Presseau, J., Hobbs, N., & Araujo-Soares, V. (2012). Testing self-regulation interventions to increase walking using factorial randomized n-of-1 trials. *Health Psychology*, 31(6), 733–737. doi:10.1037/a0027337
- Sonnier, B. L. (2002). *Effects of self-monitoring and monetary reward on fluid adherence among adult hemodialysis patients* (Unpublished PhD dissertation). University of North Texas.
- Soroudi, N., Perez, G. K., Gonzalez, J. S., Greer, J. A., Pollack, M. H., Otto, M. W., & Safren, S. A. (2008). CBT for medication adherence and depression (CBT-AD) in HIV-infected patients receiving methadone maintenance therapy. *Cognitive and Behavioral Practice*, 15(1), 93–106. doi:10.1016/j.cbpra.2006.11.004
- Stone, A. A. (2000). *The science of self-report: Implications for research and practice*. Mahwah, NJ: Lawrence Erlbaum.
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Pearson/A&B.

- Tate, R. L., Perdices, M., Rosenkoetter, U., Wakim, D., Godbee, K., Togher, L., & McDonald, S. (2013). Revision of a method quality rating scale for single-case experimental designs and n-of-1 trials: The 15-item risk of bias in n-of-1 trials (RoBiNT) scale. *Neuropsychological Rehabilitation*, 23(5), 619–638. doi:10.1080/09602011.2013.824383
- Tate, R. L., Rosenkoetter, U., Vohra, S., Horner, R., Kratochwill, T., Sampson, M., ... Wilson, B. (2016). The single-case reporting guideline in behavioural interventions (SCRIBE) 2016 statement. *Journal of Clinical Epidemiology*, 73, 142–152. doi:10.1016/j.jclinepi.2016.04.006
- Vail-Gandolfo, N. (2009). *The effects of a motivational program on glycemic control, physical activity and weight loss in individuals with type 2 diabetes mellitus (Unpublished PhD dissertation)*. Hofstra University, New York.
- Vohra, S., Shamseer, L., Sampson, M., Bukutu, C., Schmid, C. H., Tate, R., ... Moher, D. (2015). Consort extension for reporting n-of-1 trials (CENT) 2015 statement. *Journal of Clinical Epidemiology*. doi:10.1016/j.jclinepi.2015.05.004
- Ward, C. M., Brinkman, T., Slifer, K. J., & Paranjape, S. M. (2010). Using behavioral interventions to assist with routine procedures in children with cystic fibrosis. *Journal of Cystic Fibrosis*, 9(2), 150–153. doi:10.1016/j.jcf.2009.11.008
- Wright, N. R., & Thompson, C. (2002). Withdrawal from alcohol using monitored alcohol consumption: A case report. *Alcohol and Alcoholism*, 37(4), 344–346. doi:10.1093/alcalc/37.4.344
- Yngman-Uhlin, P., Fernstrom, A., Borjeson, S., & Edell-Gustafsson, U. (2012). Evaluation of an individual sleep intervention programme in people undergoing peritoneal dialysis treatment. *Journal of Clinical Nursing*, 21(23–24), 3402–3417. doi:10.1111/j.1365-2702.2012.04282.x